Supporting Information

Intermolecular Hydrogen Bonding Delineates Stability of Non-canonical Adenine Base Pairs: A First-Principles Study

Nicholas Adu-Effah and Nabanita Saikia[*]

Department of Chemistry, New Mexico Highlands University, Las Vegas, NM 87701, USA

Corresponding author:

*E-mail: nsaikia@nmhu.edu

ORCID

Nabanita Saikia: 0000-0002-0684-260X

Table S1. Geometric properties of A·A dimers stabilized by N-H…N and C-H…N hydrogen bonds in the gas and solvent phase at the PBE-D2/6-31G(d,p) level. D-H…A represents the intermolecular hydrogen bond between donor (D-H) and acceptor (A) sites.

Dimer	Intermolecular hydrogen bonded atoms	Gas Phase		Solvent Phase	
		D-H…A (Å)	Bond angle (°)	D-H…A (Å)	Bond angle (°)
AA_fHS_(79)(83)	N7-H…N9	2.84	161	2.80	163
	C8-H…N3	3.18	147	3.28	140
AA_mSS_(23)(39)	C2-H…N3	3.38	151	3.46	147
	N9-H…N3	2.84	165	2.80	166
AA_mWH_(18)(67)	C8-H…N1	3.17	143	3.23	136
	N6-H…N7	2.93	170	2.91	170
AA_fWW_(16)(21)	N6-H…N1	2.94	174	2.91	175
	C1-H…N2	3.34	148	3.38	146
AA_mWS_(12)(63)	C2-H…N1	3.34	147	3.39	142
	N6-H…N3	2.93	173	2.90	173
AA_fWS_(63)(72)	N6-H…N3	2.95	172	2.91	173
	C2-H…N7	3.32	151	3.35	144

Table S2. Geometric properties of A·A dimers stabilized solely by C-H…N hydrogen bonds in gas and solvent phase at the PBE-D2/6-31G(d,p) level of theory. D-H…A represents the intermolecular hydrogen bond between donor (D-H) and acceptor (A) sites.

Dimer	Intermolecular hydrogen bonded atoms	Gas Phase		Solvent Phase	
		D-H…A (Å)	Bond angle (°)	D-H…A (Å)	Bond angle (°)
AA_mHH_(78)(87)	C8-H…N7	3.23	142	3.25	139
	C8-H…N7	3.23	142	3.25	139
AA_fWH_(18)(27)	C8-H…N1	3.24	144	3.24	145
	C2-H…N7	3.40	147	3.45	146
AA_mHS_(72)(83)	C2-H…N7	3.39	146	3.45	144
	C8-H…N3	3.23	144	3.23	144
AA_mWW_(12)(21)	C2-H…N1	3.38	146	3.42	148
	C2-H…N1	3.38	146	3.42	148
AA_mSS_(23)(32)	C2-H…N3	3.37	144	3.41	147
	C2-H…N3	3.37	144	3.41	147

Table S3. BSSE (basis set superposition error) corrected binding energy (in units of kcal/mol) of N-H…N stabilized A·A dimers in gas phase. All calculations are performed using 6-31G(d,p) basis set.

A·A dimers	PBE-D2	wB97xD-D2	M06-2X	B97-D3	B3LYP
AA_fWS_(19)(63)	-19.79	-14.17	-12.24	-16.93	1.98
AA_mHH_(67)(76)	-13.96	-9.02	-6.92	-11.53	6.89
AA_mHS_(63)(79)	-18.19	-12.83	-10.53	-15.38	3.04
AA_mSS_(39)(93)	-23.10	-17.50	-15.36	-19.98	-1.22
AA_mWW_(16)(61)	-17.00	-11.47	-9.67	-14.41	4.52
AA-fWH_(16)(67)	-15.68	-10.45	-8.47	-13.14	5.45

A·A dimers	CAM-B3LYP		PBE1PBE		HSEH1PBE
	D3	No D3	D3	No D3	
AA_fWS_(19)(63)	-18.04	-15.27	-18.38	-15.65	-15.95
AA_mHH_(67)(76)	-12.68	-10.11	-13.01	-10.26	-10.67
AA_mHS_(63)(79)	-16.63	-13.99	-17.05	-14.34	-14.66
AA_mSS_(39)(93)	-21.39	-18.56	-21.77	-18.97	-19.16
AA_mWW_(16)(61)	-15.25	-12.60	-15.59	-12.97	-13.28
AA-fWH_(16)(67)	-14.19	-11.59	-14.51	-11.86	-12.18

Table S4. The gas phase binding energy (E_b) of A·A dimers. The optimized values correspond to the M06-2X/6-31G(d,p) level.

A·A dimers	E_b (kcal/mol)
AA_mSS_(39)(93)	-20.61
AA_fWS_(19)(63)	-17.58
AA_mHS_(63)(79)	-15.92
AA_mWW_(16)(61)	-15.10
AA-fWH_(16)(67)	-13.90
AA_mHH_(67)(76)	-12.49
AA_fHS_(79)(83)	-12.74
AA_mSS_(23)(39)	-10.81
AA_mWH_(18)(67)	-10.32
AA_fWW_(16)(21)	-8.67
AA_mWS_(12)(63)	-8.52
AA_fWS_(63)(72)	-8.03
AA_mHH_(78)(87)	-6.44
AA_fWH_(18)(27)	-5.25
AA_mHS_(72)(83)	-5.14
AA_mWW_(12)(21)	-4.00
AA_mSS_(23)(32)	-3.79

Table S5. Deformation energy, E_{def} and total energy, E_{tot} (in parenthesis) in units of kcal/mol for A·A dimers in the gas phase. E_{tot} is calculated as, $E_{tot} = E_{def} + E_{b, BSSE}$,[1] where $E_{b, BSSE}$ is the BSSE corrected interaction energy. The E_{tot} is calculated for the A·A dimers stabilized by N-H…N bonds at the PBE-D2/6-31G(d,p) level.

| A·A dimers	E_{def} (E_{tot})		
	PBE-D2	M06-2X	
	6-31G(d,p)	6-311++G(d,p)	6-311++G(2d,2p)
N-H…N bonds			
AA_mSS_(39)(93)	3.30 (-19.80)	1.31 (-16.13)	1.63 (-15.10)
AA_fWS_(19)(63)	3.14 (-16.64)	1.31 (-13.10)	1.60 (-12.20)
AA_mHS_(63)(79)	3.10 (-15.10)	1.39 (-11.47)	1.69 (-10.62)
AA_mWW_(16)(61)	2.36 (-14.65)	0.93 (-11.07)	1.18 (-10.26)
AA-fWH_(16)(67)	2.22 (-13.46)	0.94 (-9.98)	1.19 (-9.23)
AA_mHH_(67)(76)	2.03 (-11.93)	0.96 (-9.73)	1.14 (-8.82)
C-H…N and N-H…N bonds			
AA_fHS_(79)(83)	1.16	0.58 (-8.63)	0.69 (-8.56)
AA_mSS_(23)(39)	1.05	0.46 (-8.09)	0.66 (-7.53)
AA_mWH_(18)(67)	0.99	0.48 (-7.45)	0.60 (-7.01)
AA_fWW_(16)(21)	0.81	0.39 (-6.42)	0.52 (-5.60)
AA_mWS_(12)(63)	0.76	0.32 (-6.03)	0.54 (-5.42)
AA_fWS_(63)(72)	0.99	0.19 (-5.71)	0.38 (-5.14)
C-H…N bonds			
AA_mHH_(78)(87)	0.20	0.14 (-4.76)	0.19 (-4.65)
AA_fWH_(18)(27)	0.15	0.10 (-3.66)	0.20 (-3.46)
AA_mHS_(72)(83)	0.14	0.07 (-3.65)	0.24 (-3.37)
AA_mWW_(12)(21)	0.12	0.06 (-2.45)	0.23 (-2.17)
AA_mSS_(23)(32)	0.11	0.06 (-2.39)	0.20 (-2.18)
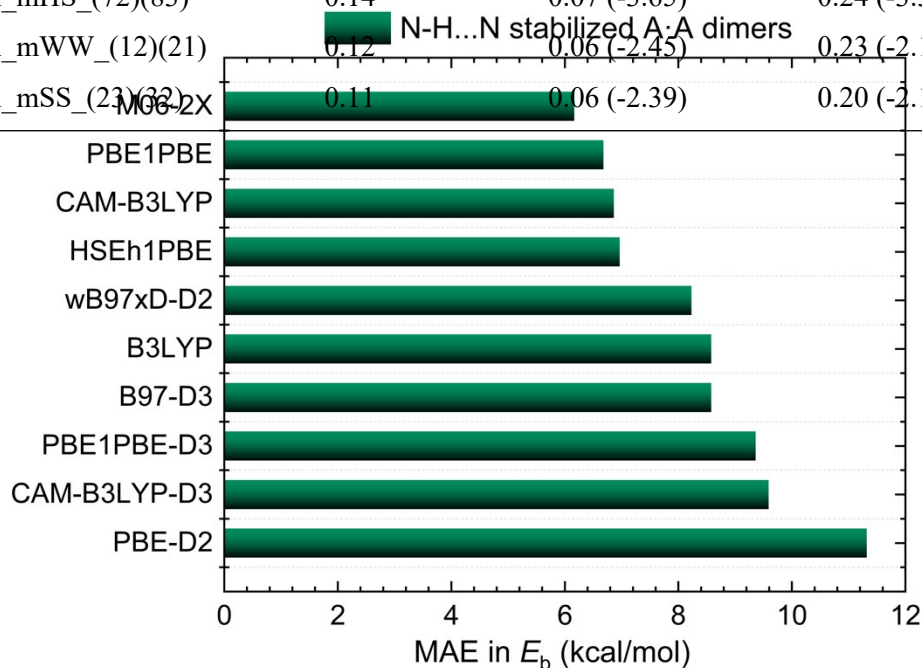
Figure S1. The MAE in E_b, for N-H…N stabilized A·A dimers in the gas phase benchmarked against MP2 level.

PBE-D2 overestimates the MAE in E_b by ~11.3 kcal/mol followed by CAM-B3LYP-D3 when benchmarked against MP2. PBE functional has been shown to overestimate hydrogen bond strengths, and with the inclusion of the D2 dispersion term, it is overcounted further. The results suggest a substantial dependence on stability of adenine base pairs on the long-range vdW interaction. Of the functionals considered, MAE in E_b was found to be lowest in the order: M06-2X (6.17) > PBE1PBE (6.68) > CAM-B3LYP (6.86) > HSEH1PBE (6.97). These four functionals may provide a better prediction of E_b for hydrogen bond stabilized A·A dimers. Based on these results, it can be concluded that the choice of functional and basis set is fundamental when considering a system of study.

HOMO–LUMO Energy Gap and Frontier Molecular Orbitals

The HOMO (Highest occupied molecular orbital) – LUMO (Lowest unoccupied molecular orbital) energy gap is used to predict molecular stability and/or reactivity of chemical species. The E_{HOMO} signifies the electron-donating ability of a molecule; with higher values indicating an increased ability to donate electrons to the unoccupied molecular orbital of a receptor. The localization of

electron density in HOMO indicates nucleophilic sites, while the localization of LUMO suggests electrophilic sites. The E_{LUMO} is associated with the molecule's ability to accept electrons, with a lower value indicating a greater ability to accept electrons. A comparison of the HOMO–LUMO energy gap of A·A dimers in the gas phase is shown in Figure S1. The energy gap is classified into three categories based on the interacting edges (N-H…N, C-H…N and N-H…N, and C-H…N).
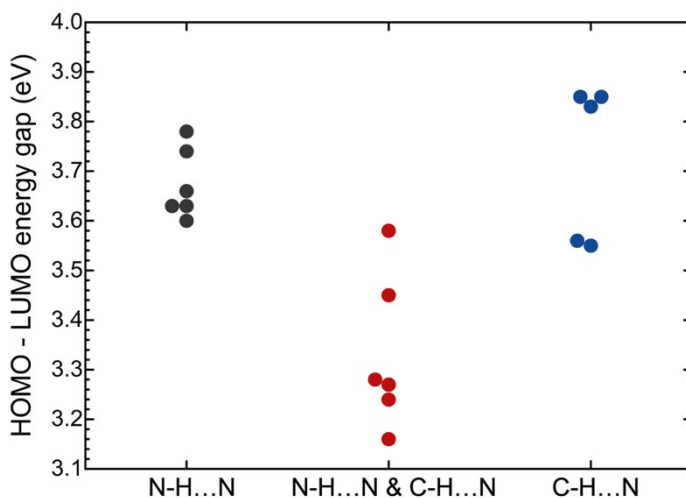

Figure S2. Comparison of the HOMO-LUMO energy gap of A·A dimers calculated in gas phase at the PBE-D2/6-31G(d,p) level of theory.

For N-H…N stabilized dimers, the HOMO–LUMO energy gap varies between 3.60 – 3.78 eV (Figure S1, grey circles) and is mostly clustered for three of the dimers. There is not a major fluctuation in the calculated energy gap of these dimers. For dimers stabilized by both C-H…N and N-H…N hydrogen bonds, the HOMO–LUMO energy gap varies between 3.16 – 3.45 eV (Figure S1, red circles). Two of the dimers have almost comparable energy gap values. In C-H…N stabilized dimers, three of the dimers have comparable energy gap values between 3.83 – 3.85 eV, while two of the dimers have an energy gap of 3.55 and 3.56 eV, respectively (Figure S1, blue circles). A comparison of the HOMO–LUMO energy gap suggests that structures stabilized by both C-H…N and N-H…N hydrogen bonds show a spread in the energy gap and have lower energy gap values than the other two categories. Presence of both C-H…N and N-H…N bonds lead to a subsequent lowering of the energy gap, which be correlated to the interacting edges and the donor-acceptor groups that are involved in hydrogen bonding. Nonetheless, a distinct trend in variation

of HOMO–LUMO energy gap based on the interacting edges involved in intermolecular hydrogen bonding could not be drawn.

Frontier Molecular Orbital

The frontier molecular orbitals play an important role in interpreting the stability and chemical reactivity of molecules, as they are based on the HOMO and LUMO isosurface. Figures S2 and S3 depict the frontier molecular orbitals corresponding to the HOMO and LUMO of N-H…N stabilized adenine dimers. Here, only the N-H…N bonded dimers are considered. In AA_mSS_(39)(93), HOMO and LUMO is distributed along the adenine dimer. The HOMO is delocalized along the aromatic ring with contributions on the –NH$_2$ groups while the LUMO is somewhat localized on the atoms of the ring (Figure S2). In AA_fWS_(19)(63), frontier orbitals corresponding to HOMO is delocalized on one of the adenine molecules while the LUMO is distributed on the other adenine molecule. In AA_mHS_(63)(79), HOMO is delocalized along the dimer while LUMO is localized on the atoms of one of the adenine molecules.
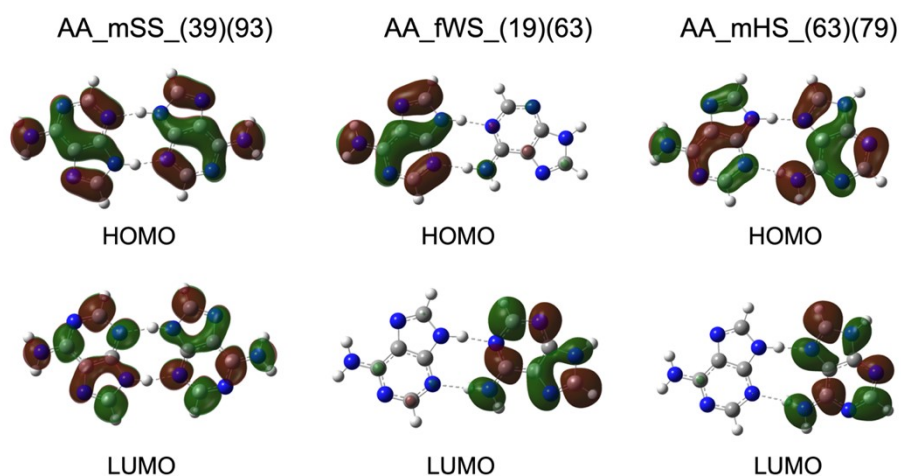

Figure S3. The HOMO and LUMO isosurface of A·A dimers stabilized by N-H…N hydrogen bonds.
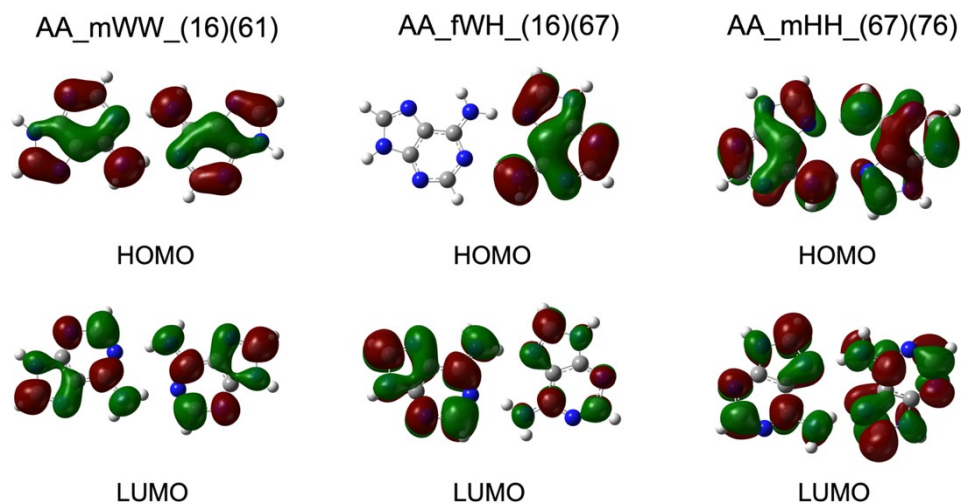
Figure S4. The HOMO and LUMO isosurface of A·A dimers stabilized by N-H…N hydrogen bonds.

In AA_mWW_(16)(61), both HOMO and LUMO are distributed along the dimer, HOMO is delocalized along the aromatic rings while LUMO is localized on the atoms of adenine. Except for the HOMO of AA_fWH_(16)(67), which is delocalized on one of the adenine molecules, for the three conformations shown in Figure S3, the HOMO and LUMO are distributed along the dimer. In general, HOMO is delocalized (bonding in character) while LUMO is localized (antibonding).

References

1. Sharma, P.; Chawla, M.; Sharma, S.; Mitra, A. On the role of Hoogsteen:Hoogsteen interactions in RNA: ab initio investigations of structures and energies. *RNA* **2010**, *16* (5), 942-957. DOI: 10.1261/rna.1919010